

Capítulo 11

Sistemas de predicción de lenguas visogestuales basados en IA y su aplicación en la Lengua de Señas Mexicana (LSM)

*Guillermo Hernández-Santana
Irvin Hussein López-Nava*

<https://doi.org/10.61728/AE24001113>



Introducción

Es indudable que a partir del desarrollo de la inteligencia artificial (IA) y los sistemas de procesamiento de lenguaje natural (PLN), las computadoras comenzaron a comprender, interpretar y generar lenguaje humano en sus diversas formas, sin embargo, estos sistemas cuentan con un mayor desarrollo en relación con las lenguas orales que a las lenguas de modalidad visogestual. Esto responde a la colección de datos lingüísticos que contamos en la actualidad, la cual, es significativamente mayor en lenguas orales que señadas. Es decir, que el desarrollo de la IA y del PLN tiene en la actualidad avances importantes en cuanto a la capacidad de las máquinas para interpretar y generar texto en diversos contextos, desde traductores automáticos hasta asistentes virtuales, pasando por chatbots.

En este capítulo discutimos la importancia del reconocimiento automático de las lenguas de modalidad visogestual, vital para la inclusión y comunicación efectiva de las personas sordas. En primer lugar, estos sistemas representan un avance tecnológico significativo en el campo de la accesibilidad, permitiendo una mayor autonomía y participación de este sector de la población en diversos contextos sociales y profesionales. Algunos de los beneficios de estas aplicaciones y su implementación pueden repercutir en tecnologías de reconocimiento automático de lenguas de señas las cuales pueden facilitar la comunicación entre personas sordas y oyentes, reduciendo las barreras lingüísticas y promoviendo la inclusión social, entre otros beneficios.

En segundo lugar, el desarrollo de sistemas de reconocimiento de lenguas de señas impulsa la innovación en áreas como la IA. De hecho, estos sistemas requieren algoritmos complejos y sofisticados que puedan procesar y entender las expresiones faciales en tiempo real, lo que representa un desafío importante en el campo de la visión por computadora. En este sentido, la investigación en esta área contribuye al avance del conocimiento en técnicas de aprendizaje profundo, procesamiento de imágenes y análisis de movimiento, aplicables en una amplia gama de tecnologías emer-

gentes, por ejemplo, los gestos que acompañan el lenguaje oral y también aspectos relacionados a la gestualidad humana en diferentes situaciones contextuales.

Este tipo de tecnología utiliza modelos de aprendizaje automático y visión por computadora para interpretar los gestos y movimientos característicos de las lenguas de señas, convirtiéndolos en texto o voz y sirve como herramienta de traducción en medios educativos para personas sordas que necesitan desarrollar habilidades lingüísticas y para oyentes que deseen aprender lenguas de señas. En este sentido, la integración de esta tecnología en plataformas educativas y aplicaciones móviles ofrecen oportunidades para el aprendizaje interactivo y accesible de lenguas de señas, promoviendo su difusión y comprensión (Gortarez-Pelayo et al. 2023).

Los sistemas PLN, por su parte, utilizan algoritmos y modelos estadísticos para analizar y comprender el significado del lenguaje humano, incluidas las variaciones en la forma de expresión, la semántica y la gramática. Pero el estado actual del estudio de las lenguas de señas no nos ha permitido reconocer del todo cuáles son los rasgos significativos de una lengua de modalidad visogestual a diferencia del conocimiento y bancos de datos que tenemos en materia de lenguas orales, la cual es significativamente menor. Particularmente, es necesario el desarrollo y comprensión de aquellos rasgos gramaticales más allá de las configuraciones manuales, es decir, aquellos rasgos de tipo gestuales que cargan información tempoaspectual de las lenguas señadas (Fridman, 2005).

En este capítulo discutimos que los sistemas basados en IA pueden adaptarse para reconocer y procesar patrones específicos de gestos y movimientos asociados con la lengua de señas. De hecho, estos sistemas cuentan con la capacidad para entender la estructura y el significado de la lengua de señas, lo cual puede contribuir directamente a la creación de tecnologías más efectivas en la interpretación y traducción automática, promoviendo así la inclusión y accesibilidad para la comunidad de personas Sordas. En la siguiente sección analizamos la evolución de IA y cómo se ha visto involucrada la tecnología de reconocimiento de imagen desde los primeros patrones de movimiento mediante diferentes tecnologías hasta cámaras y algoritmos para reconocer los patrones de movimiento.

Desarrollo del reconocimiento de lenguas de señas mediante IA

La historia del reconocimiento automático de lenguas de señas mediante inteligencia artificial (IA) se remonta a las últimas décadas del siglo XX, cuando los avances en la visión por computadora y el procesamiento de señales digitales comenzaron a permitir la detección y el seguimiento de gestos humanos. Los primeros sistemas se basaban en técnicas rudimentarias de reconocimiento de patrones y requerían de guantes sensorizados o marcadores en las manos para detectar los movimientos. Estos sistemas, aunque innovadores, eran limitados en su capacidad para reconocer gestos complejos y a menudo eran intrusivos y poco prácticos para los usuarios (Sturman y Zeltzer, 1994). La dependencia de hardware especializado y la falta de portabilidad eran desafíos significativos en esta fase inicial.

Con la llegada del siglo XXI, la visión por computadora experimentó avances significativos, impulsados en parte por el aumento en la capacidad de procesamiento y algoritmos más sofisticados. Esto permitió el desarrollo de sistemas de reconocimiento de lenguas de señas más avanzados, que podían capturar gestos mediante cámaras estándar sin la necesidad de dispositivos adicionales. La aplicación de algoritmos de aprendizaje automático comenzó a permitir un reconocimiento más preciso de los gestos y una mejor interpretación del contexto de las señas (Murakami y Taguchi, 1991). Sin embargo, estos sistemas aún luchaban con la variabilidad en la ejecución de las señas entre diferentes usuarios y las limitaciones en la diferenciación de gestos similares.

La introducción de las redes neuronales convolucionales (CNN) y las redes neuronales recurrentes (RNN) en el campo del reconocimiento de lenguas de señas marcó un hito importante. Estas tecnologías permitieron a los investigadores abordar las complejidades inherentes al lenguaje de señas, como la variabilidad en la ejecución de las señas y la importancia del contexto y la fluidez en la señalización. La capacidad de las CNN para procesar imágenes y de las RNN para entender secuencias temporales resultó ser particularmente útil para interpretar los movimientos dinámicos y las expresiones faciales asociadas con las lenguas de señas (Pigou et al., 2015). Este avance representó un paso significativo hacia sistemas más robustos y precisos.

A pesar de los avances tecnológicos, el reconocimiento de lenguas de señas por IA continuó enfrentando desafíos significativos. Uno de los más importantes era la variabilidad en la forma en que diferentes personas realizan las señas, lo que se veía agravado por factores como la edad, el género y el origen cultural; sin mencionar la gran cantidad de lenguas de señas que existen a nivel mundial. Además, el contexto en el que se realiza una seña puede cambiar su significado, lo que presenta un desafío adicional para su interpretación correcta. En la actualidad sabemos que las lenguas de señas son ricas en matices y requieren un nivel de comprensión contextual que los sistemas de IA inicialmente lucharon por alcanzar (Koller et al., 2015).

La recolección de grandes conjuntos de datos en determinadas lenguas de señas y el avance en técnicas de aprendizaje profundo, adoptadas principalmente de otras áreas, ha permitido entrenar modelos de IA más robustos y precisos. El aprendizaje profundo, particularmente con arquitecturas de redes neuronales profundas, ha sido fundamental para capturar las complejidades y sutilezas de las lenguas de señas, mejorando significativamente la precisión del reconocimiento (Camgoz et al., 2018; 2017).

La integración de diferentes tecnologías de sensado, como el Microsoft Kinect o el Leap Motion, junto con avances en realidad aumentada, ha abierto nuevas posibilidades para el reconocimiento de lenguas de señas. Estos dispositivos proporcionan una captura de movimiento más precisa y menos intrusiva, lo que permite una detección detallada de los gestos y expresiones faciales. Además, la realidad aumentada ofrece formas innovadoras de interactuar con los sistemas de reconocimiento de lenguas de señas, mejorando la experiencia del usuario y facilitando la comunicación en tiempo real (Lang et al., 2020).

El progreso en el reconocimiento de lenguas de señas por IA ha llevado a aplicaciones prácticas que mejoran la accesibilidad para las personas sordas. Estas aplicaciones incluyen sistemas de traducción en tiempo real, asistentes virtuales personales y herramientas educativas. Estos sistemas no solo proporcionan medios para una comunicación más efectiva, sino que también promueven la inclusión y la igualdad de acceso a servicios e información para la comunidad Sorda (Bragg et al., 2019). En la siguiente sección hacemos una revisión del estado del arte en los sistemas de predicción de lenguas de señas.

Sistemas de predicción en diferentes lenguas de señas en el mundo

Como se señaló en la sección anterior, dos aspectos a considerar en el conocimiento actual de las lenguas de señas son la variación del habla individual, así como de las formas y expresiones que se desarrollaron de manera independiente en diferentes contextos geográficos, es decir, tenemos poca información acerca de la manera en que funciona la gramática y el discurso de los idiomas visogestuales. Aunado a esto, no podemos esperar uniformidad de las expresiones en las diferentes lenguas de señas, por el contrario, lo esperable es encontrar variación lingüística, tanto en el eje diacrónico como en el geográfico.

Lo anterior se explica por algunos aspectos de las lenguas naturales. Estos son la variación, la distancia geográfica, así como el desarrollo de neologismos, lo cual es indudable en situaciones de contacto lingüístico y ante la necesidad de crear léxico ante nuevas situaciones comunicativas.

Por un lado, el conocimiento actual de patrones, expresiones y reglas de formación de las lenguas de señas en la actualidad es limitado, pues no contamos con tantas bases de datos o corpora lingüística que nos informe cuáles son los patrones articulatorios que conforman la gramática de la lengua de señas y cuáles no; es decir, criterios que aporten información de las diferencias estructurales y pragmáticas de la lengua de señas que alimenten modelos de reconocimiento de expresiones mediante IA.

En este apartado revisamos una variedad de modelos que se han llevado a cabo para el reconocimiento de patrones lingüísticos de lenguas de modalidad visogestual en la actualidad. Para el desarrollo de este apartado, se llevó a cabo una búsqueda de literatura en las plataformas Scopus, Google Académico y la base de datos de educación ERIC. La estrategia de búsqueda en estas plataformas se definió mediante la combinación de términos específicos: “lengua de señas mexicana, inteligencia artificial, reconocimiento de datos, gestualidad, neural language processing” de artículos realizados a partir de 2023 considerando el rápido auge que han tenido los sistemas de procesamiento de lenguas naturales a partir del desarrollo de los chatbots. Se consideraron tanto trabajos escritos en español como en inglés. Para garantizar la consistencia en la selección de estudios, se exclu-

yeron aquellos relacionados con temas de salud y también se excluyeron trabajos de tesis.

Este procedimiento aseguró una recopilación exhaustiva y focalizada de la literatura pertinente mediante el metaanálisis PRISMA (Moher et al., 2015), facilitando así un análisis detallado de la intersección entre la Lengua de Señas Mexicana y la inteligencia artificial en el contexto del reconocimiento de datos y gestualidad. Además, se aplicó un criterio de exclusión al descartar los artículos que no fueron revisados por pares, asegurando la calidad y confiabilidad de la información recopilada. A partir de esta revisión se realizó un análisis cualitativo de artículos referidos en las obras para ampliar la información al respecto que nos pudiera dar información acerca de los diferentes patrones articulatorios que deberían considerarse en los idiomas de modalidad visogestual, así como aquellos aspectos relevantes desde un punto de vista gramatical. Adicionalmente, se incluyeron algunas búsquedas particulares de tecnología que había sido reportada en las obras pero que no necesariamente estaban incluidas en las bases de datos.

Para llevar a cabo la búsqueda en lengua de señas mexicana, se optó por metodología cualitativa que diera cuenta de las tecnologías relacionadas con este idioma en diversas fuentes y portales especializados. La metodología incluyó la revisión y consulta de plataformas académicas (Scopus, ProQuest e IEEE Xplore) con términos como “lengua de señas mexicana”, “inteligencia artificial”, “reconocimiento de datos”, “gestualidad” y “procesamiento de lenguaje natural” y también optamos por los antecedentes que los autores tienen en el conocimiento de la bibliografía especializada en la lengua de señas mexicana, dado que uno de ellos ha desarrollado un sistema de reconocimiento de patrones lingüísticos (Morfín-Chávez, Gortarez-Pelayo y Lopez-Nava, 2023). Este enfoque metodológico permitió obtener una visión integral de las tecnologías asociadas con la lengua de señas mexicana.

En el ámbito de la captura del lenguaje de señas, se busca registrar gestos utilizando configuraciones de sensores apropiadas con el fin de obtener información discriminativa de los elementos que se capturen. Este proceso facilita el estudio, reconocimiento y representación 3D de las señas en etapas posteriores, permitiendo también la construcción de

conjuntos de datos extensos. La variedad de configuraciones de sensores empleados destaca la adaptabilidad de diversas aplicaciones (Papastratis, 2021). En los siguientes apartados, se presentan los dispositivos de captura más importantes que se han empleado en los últimos años en materia de lengua de señas.

Sensores de visión 2D

Los avances en las tecnologías de sensado han marcado un progreso notable, con categorías diferenciadas según la tecnología empleada y los datos capturados. En el año 2023, Morfín-Chávez y su equipo implementaron una estrategia innovadora utilizando una cámara 2D junto con la biblioteca MediaPipe para realizar la segmentación de la mano y extraer los puntos clave (keypoints) en dos dimensiones, y aproximando la tercera dimensión. Este enfoque se destaca por su simplicidad y accesibilidad, ya que permite la captura de imágenes desde dispositivos comunes, como cámaras web o teléfonos móviles, y la extracción de características detalladas de las manos mediante algoritmos avanzados de visión por computadora. La simplicidad de este método no solo facilita la portabilidad y respuesta inmediata, sino que también se trata de una solución idónea para aplicaciones que requieren facilidad de uso, como aplicaciones móviles o sistemas basados en la web. Esta combinación de una cámara 2D estándar y la biblioteca MediaPipe representa una contribución significativa al campo del reconocimiento de lenguas de señas, al ofrecer una alternativa para la captura y análisis de elementos gestuales, allanando el camino para soluciones prácticas y generalizadas en el ámbito de la tecnología de asistencia.

Sensores de visión 3D

Entre otros inputs de imagen que se han utilizado se reporta el uso del sensor Kinect como medio captura de datos de profundidad y de color RGB (red, green, blue) así como de la estructura esquelética del individuo. Se han encontrado aplicaciones similares con cámaras de visión artificial. El uso de este sensor destaca de una categoría avanzada de dispositivos de sensado ya que está compuesto tanto de una cámara 2D tradicional

como un sensor de profundidad (García-Bautista et al., 2017; Sosa-Jiménez et al., 2017). Por lo que el sensor Kinect proporciona capacidades de seguimiento a partir de keypoints, lo que lo hace ideal para aplicaciones que requieren un alto nivel de detalle y precisión en el reconocimiento de gestos. Esta tecnología es capaz de capturar no solo el movimiento de las manos, sino también la postura y los movimientos del cuerpo, ofreciendo una comprensión más completa de las expresiones que puede registrar. Este sensor es especialmente útil en entornos de investigación y aplicaciones que requieren una alta fidelidad en la captura de datos gestuales.

Otros sensores que se han utilizado son Leap Motion, que tiene como característica poder capturar posiciones 3D de la mano y los dedos, para grabar gestos y palabras de lenguas de modalidad visogestual. Sin embargo, una de las desventajas de este tipo de sensor es que es necesario operar cerca del sujeto, aunque una de sus ventajas es su bajo requerimiento computacional, el cual les permite su uso en aplicaciones en tiempo real y es capaz de reconocer las posiciones 3D de manos y dedos.

Sensores fisiológicos

En el ámbito del reconocimiento de las lenguas de señas, otra tecnología es la utilización de sensores fisiológicos relacionados con la captura del movimiento muscular (Papastratis, 2021). Esta tecnología permite explorar nuevas formas de capturar y comprender las configuraciones manuales utilizadas en la lengua de señas. Un ejemplo paradigmático es el brazalete Thalmic MYO, reconocido por su capacidad para capturar actividades electromiográficas con precisión.

Este dispositivo ha sido específicamente empleado para el reconocimiento del alfabeto en la lengua de señas irlandesa pues su diseño innovador posibilita la captura de señales electromiográficas, proporcionando una visión detallada de los movimientos y gestos manuales asociados con la comunicación en lengua de señas. Al utilizar tecnología RFID, el brazalete Thalmic MYO ofrece un enfoque no intrusivo para el reconocimiento, ya que puede capturar datos de manera continua y en tiempo real. En este sentido, la incorporación de sensores de movimiento, como el Thalmic MYO, no solo amplía las posibilidades de reconocimiento de lenguas de

señas, sino que también abre la puerta a aplicaciones más avanzadas en el ámbito de la tecnología de asistencia.

Guantes de datos y otros dispositivos

En el ámbito del reconocimiento de lenguas de señas, los guantes representan otra tecnología de precisión en tiempo real para traducir señas. Tenemos en caso del denominado BrithtSign (Ompi, 2019) que tiene la capacidad para capturar los movimientos detallados de los articuladores manuales para proporcionar un nivel de detalle excepcional en el reconocimiento de datos. Aunque ofrecen una precisión importante, ajustar los componentes de estos guantes puede resultar en un proceso laborioso. Esto puede implicar una curva de aprendizaje para los usuarios y, en algunos casos, limitar su adopción, especialmente en situaciones reales.

Adicionalmente, se ha explorado el uso de dispositivos como los relojes inteligentes (Herrera Crespo, 2021). Estos dispositivos ofrecen una configuración más sencilla en comparación con los guantes de datos. Algunos enfoques han implementado relojes inteligentes para recopilar señales de señas, que luego se transmiten a dispositivos móviles para su procesamiento y traducción en tiempo real. Aunque esta configuración simplificada puede ser más accesible, es importante considerar las limitaciones en términos de la cantidad de datos que pueden reconocer (Papastratis, 2021).

Como se ha podido observar, existen diferentes tecnologías que se han implementado para el reconocimiento de señas, destacan las cámaras 2D, sin embargo, también podemos ver guantes, sensores y relojes inteligentes. Cada proyecto de reconocimiento tiene sus propios retos y estrategias de sensado. Estas tecnologías ponen énfasis en el reconocimiento de las configuraciones manuales y rasgos reconocibles en la dactilología, el movimiento del cuerpo, y otros también destacan algunos rasgos adicionales, como el inicio y fin de una seña y la posición de la mano. Si bien la elección suele depender de las necesidades particulares de la aplicación y de las preferencias de los usuarios finales, estos desarrollos no solo amplían las opciones tecnológicas disponibles, sino que también subrayan la importancia de equilibrar la precisión con la usabilidad en el diseño de soluciones para la comunidad de usuarios de lenguas de señas. Otro punto por considerar

es que estas tecnologías se han desarrollado en una cantidad limitada de lenguas de señas en el mundo. Esto implica un desequilibrio en aquellos contextos menos desarrollados, por ejemplo, países subdesarrollados, que no cuentan con tantos recursos, tecnología o buenas políticas y prácticas educativas en materia de inclusión. En la siguiente sección, se ofrece un panorama de los principales sistemas visogestuales en los que se ha diseñado un sistema de sensado de patrones lingüísticos, mostrando así una generalización de los avances tecnológicos en este campo.

Evolución y reconocimiento de patrones lingüísticos a nivel global

A medida que la tecnología ha avanzado, los sistemas de sensado y reconocimiento aplicados a las lenguas de señas han experimentado notables desarrollos. Sin embargo, es importante señalar que estos avances se han enfocado predominantemente en un conjunto limitado de lenguas de señas a nivel mundial. Entre las lenguas de señas que han sido objeto de estudio y desarrollo de sistemas tecnológicos, destacan la alemana, de Hong Kong (en China), griega, británica, americana, árabe e india (Papastratis, 2021). En esta sección, exploramos algunas de las tecnologías específicas que se han aplicado en el reconocimiento de patrones lingüísticos a partir del desarrollo de la IA en estas lenguas, así como estudios relevantes que han contribuido a la evolución de estos sistemas. Como se verá, estas lenguas representan tan solo una fracción de la diversidad lingüística presente en las diversas comunidades sordas alrededor del mundo.

A nivel global, se han registrado de 130 a 271 lenguas de señas. En este ámbito, los sistemas de sensado y reconocimientos aplicados a las lenguas de señas han experimentado desarrollos relevantes, pero es importante destacar que estos avances se han centrado en un número limitado de lenguas de señas. Algunas de las tecnologías que se han aplicado, así como algunos estudios relevantes, se muestran en las siguientes secciones.

Lengua de Señas Indio (ISL)

Entre las diversas aplicaciones que merecen destacarse, se encuentra un estudio que se centra en el sistema de interpretación de la Lengua de Señas India (Sreemathy et al., 2023). Este sistema, diseñado para la interacción automática de signos, utiliza enfoques de aprendizaje profundo y se implementa en MATLAB como plataforma de simulación. Este integra redes preentrenadas como AlexNet, GoogleNet, VGG-16 y VGG-19 y logra precisiones notables alcanzando el 99.11 %. En este mismo sistema seña-do también encontramos un trabajo de Sugandhi (2021). Este se basa en una captura de imágenes basado en una Evaluación Bilingüe de Referencia (BLEU) para las oraciones y una evaluación manual de palabras y oraciones. Dicho sistema cuenta con un corpus rico de las 800 oraciones más comúnmente utilizadas en el idioma inglés y se toma en consideración una salida basada en la gramática de ISL que compara el output con el puntaje.

Justamente, el puntaje BLEU se calcula utilizando métodos estadísticos y de procesamiento de lenguaje natural para determinar la similitud entre las oraciones generadas por el sistema y las oraciones de referencia creadas por los expertos en ISL. Los resultados de estas evaluaciones proporcionaron una medida clara de la eficacia del sistema propuesto en la generación de SL y respaldaron su utilidad como herramienta para la comunicación efectiva entre personas con pérdida de audición y aquellos que no tienen conocimiento de SL.

Además, en un estudio reciente realizado por Kothadiya et al. (2022), se utilizaron las redes neuronales LSTM y GRU mediante una colección de gestos manuales integrando expresiones como “hola”, “buenos días” y “trabajo”. En este estudio se mostró que al aumentar el número de capas en las redes neuronales y aplicar una combinación se suele mejorar la precisión del modelo.

Lengua de Señas Americana (ASL)

En Lengua de Señas Americana, encontramos el trabajo de Schioppo y colaboradores (2019), el cual reconoce la lengua de señas mediante un auri-cular de realidad virtual para crear un entorno inmersivo. Este estudio se

basa en un controlador de Leap Motion, cuyas características se combinan con un banco aleatorio para la clasificación en tiempo real de los gestos del usuario. Para evaluar la eficacia de este enfoque, la tecnología se basó en pruebas con las 26 letras del alfabeto de la ASL en un entorno virtual y uno de los objetivos fue el aprendizaje de la lengua de señas.

En materia de aprendizaje profundo, Agarwal et al. (2021) desarrollaron DeepSign, una herramienta de software la cual es capaz de evaluar el progreso que tiene un estudiante en la alfabetización. Esta innovadora solución automatiza la puntuación de medidas de nombramiento de imágenes en lenguaje de señas, ofreciendo retroalimentación inmediata y reduciendo la carga laboral de los docentes. En esta misma lengua de señas, encontramos un método de enseñanza que está dirigido a sordos y padres oyentes mediante entornos inmersivos de realidad virtual (Quandt, 2020). En dicho mundo virtual se emplean avatares que señan y enseñan ASL de manera inmersiva. Las contribuciones significativas incluyen el uso de avatares, la integración de gestos, vocabulario básico en ASL introductorio en un entorno 3D.

Lengua de Señas Británica (BSL)

En el contexto británico, se han desarrollado diversas herramientas automáticas que reconocen datos de trayectoria de la mano, datos faciales y datos de distribución del codo para el reconocimiento de rasgos lingüísticos en la British Sign Language (BSL). Estos datos son procesados mediante un modelo de red CNN, tal como se describe en el trabajo de Liang et al. (2020). Estos esfuerzos reflejan una iniciativa más amplia en el Reino Unido, donde se promueve activamente la mejora de la accesibilidad y la inclusión, como se discute en el estudio de Gellai (2023). Además, en el ámbito de la tecnología aplicada a las lenguas de señas, se han desarrollado dispositivos como guantes capaces de reconocer las configuraciones manuales de la BSL, como se evidencia en el trabajo de Ompi (2019).

También en el ámbito de la BSL, Albanie et al. (2020) desarrollaron un enfoque para la recolección de datos destinados al reconocimiento automatizado de patrones lingüísticos mediante la producción de subtítulos alineados a través del reconocimiento de palabras clave, lograron localizar

automáticamente instancias de señas en videos continuos, lo que resultó en la creación de un conjunto de datos de un banco de mil señas. Además, mostraron que este conjunto de datos puede ser utilizado para entrenar modelos sólidos de reconocimiento de signos.

Lengua de señas Árabe (ArSL)

Otro avance significativo en el reconocimiento de gestos lo vemos en el proyecto denominado DeepArSLR (Aly y Aly, 2020). Este avance de censado en ARSL se basa en el reconocimiento de gestos aislados e independientes del señante el cual se combina una red de segmentación semántica, una red convolucional y una red LSTM bidireccional profunda. Esta dinámica permite una representación de la forma de la mano capaz de reconocer configuraciones de mano a través del modelo SOM convolucional de una sola capa. Esta tecnología también incluye una nueva red de clasificación profunda, compuesta por tres capas BiLSTM bidireccionales conectadas para el sensado de secuencias de gestos.

En ArSL también reconocemos el trabajo de Elatawy (2020) que presenta una novedosa aproximación al reconocimiento del alfabeto de la lengua de señas árabe mediante el uso de predicciones neutrosóficas y significados difusos, una propuesta que convierte la información visogestual en un dominio neutrosófico mediante un filtro gaussiano en las imágenes que captura.

Lengua de Señas Alemana (DGS)

En el contexto germánico, se ha desarrollado desde hace varios años una amplia corpora lingüística que se denomina RWTH-PHOENIX-Weather. Esta es una colección de datos extraída de la televisión pública alemana, y anotados manualmente con glosas a nivel de oración. Además, se han integrado transcripciones semiautomáticas del alemán hablado utilizando el sistema de reconocimiento de voz. Adicionalmente, se han incluido anotaciones referentes a la forma y orientación de la mano dominante para más de cuarenta mil cuadros de video, convirtiéndolo en uno de los mayores corpus que permiten la evaluación de algoritmos de seguimiento

de objetos. Asimismo, se han anotado más de dos mil signos utilizando el sistema de anotación SignWriting, centrándose en la forma, orientación, movimiento y contactos espaciales de ambas manos (Forster, 2014; Koller et al. 2015).

En la revisión de esta sección, se observa un mayor desarrollo en el reconocimiento de tecnologías aplicadas en algunas lenguas de señas en comparación con otras. Las lenguas de señas indio (ISL), americana (ASL), británica (BSL), alemana (DGS) así como la (ArSL) han sido objeto de una considerable cantidad de investigación y desarrollo, lo que ha llevado a la creación de diversos conjuntos de datos, herramientas y modelos de reconocimiento de señas específicas. Sin embargo, se nota una brecha en el desarrollo de tecnologías para lenguas de señas menos estudiadas, tal es el caso de algunas lenguas visogestuales de la región Centro y Sudamericana así como las lenguas de señas en África y Oceanía.

Es probable que el limitado desarrollo en el reconocimiento de lenguas de señas se deba en parte al hecho de que algunas lenguas de modalidad visogestual son relativamente jóvenes y el tiempo de desarrollo de tecnologías en el área no ha sido suficiente. Un ejemplo de esto es la lengua de señas de Guatemala, la cual surgió en la década de los ochenta como resultado de una política de inclusión gubernamental. Por otro lado, en países donde las políticas de inclusión han sido menos efectivas, encontramos menos reportes de sistemas de reconocimiento mediante IA. Estas disparidades subrayan la urgente necesidad de dirigir una mayor atención y recursos hacia el desarrollo de tecnologías de reconocimiento de lenguas de señas para comunidades menos representadas. Además, es crucial continuar desarrollando amplios bancos de datos lingüísticos para el desarrollo de aplicaciones que puedan servir para aplicaciones de reconocimiento lingüístico.

Reconocimiento automático en Lengua de Señas Mexicana (LSM)

Los sistemas de procesamiento de lenguaje natural pueden adaptarse para reconocer y procesar patrones específicos de gestos y movimientos asociados con una lengua de señas. De hecho, estos sistemas en la actualidad

cuentan con la capacidad para entender la estructura y el significado de la lengua de señas, lo cual puede contribuir directamente a la creación de tecnologías más efectivas en la interpretación y traducción automática, promoviendo así la inclusión y accesibilidad para la comunidad de personas Sordas.

El reconocimiento automático de la Lengua de Señas Mexicana (LSM) también ha experimentado avances notables a lo largo de los años, reflejados en varios estudios significativos. En uno de los primeros estudios identificados (Solís et al. 2014) se enfocó en el reconocimiento de 24 señas estáticas de LSM utilizando una cámara digital bajo condiciones de luz incoherente y Redes Neuronales Artificiales (ANN). Este estudio logró una alta tasa de reconocimiento, demostrando la eficacia de las ANN en la interpretación de señas de LSM.

Posteriormente, en 2015, Galicia et al. abordaron la conversión de LSM a letras del alfabeto español oral utilizando sensores de movimiento y algoritmos de clasificación avanzados, en el que alcanzaron una precisión notable en la traducción de LSM. En el mismo año, Solís et al. se centraron en el reconocimiento de señas estáticas en LSM usando Jacobi-Fourier Moments y ANN, alcanzando una alta tasa de reconocimiento. En 2016, Solís et al. se enfocaron en 21 señas estáticas de LSM, utilizando una ANN para el análisis de datos obtenidos mediante una cámara digital. Este estudio logró una tasa de reconocimiento de casi 100%, reforzando la viabilidad de utilizar aprendizaje profundo y momentos normalizados para el reconocimiento de LSM.

García-Bautista et al., propusieron en el 2017 un método para reconocer una lista básica de palabras en LSM utilizando un sensor Kinect y el algoritmo de Dynamic Time Warping, logrando una precisión media del 99.1 %. Este enfoque demostró ser altamente efectivo para el reconocimiento en tiempo real de LSM. En el mismo año, Sosa-Jiménez et al. presentaron un estudio en el que clasificaban un conjunto de señas de LSM con datos obtenidos de un Kinect y logrando una sensibilidad promedio del 86 % y una especificidad promedio del 80 %. En 2022, Sosa-Jiménez et al. desarrollaron un sistema bidireccional para LSM en servicios de salud. Utilizando un sensor Kinect y modelos ocultos de Markov (HMMs), el sistema logró tasas altas de precisión, demostrando el potencial de los sistemas de traducción bidireccional en contextos de atención médica.

El estudio de Morfín-Chávez et al., presentado en 2023, se centró en el reconocimiento del alfabeto manual de la LSM, utilizando técnicas de aprendizaje automático como Support Vector Machines. Los datos fueron adquiridos con una cámara de video convencional, y se extrajeron 21 puntos clave de las manos con la biblioteca MediaPipe para entrenar los modelos de clasificación. Los resultados al clasificar las 21 letras superaron una F1-score de 0.98. Finalmente, en 2022, Mejía-Peréz et al. introdujeron un sistema para reconocer LSM basado en gestos múltiples, utilizando una cámara de profundidad (OAK-D) y redes neuronales recurrentes para la clasificación, alcanzando una precisión del 97 % en datos limpios y 90 % en datos ruidosos. Este enfoque avanzado subraya la efectividad de combinar cámaras de profundidad y RNN para un reconocimiento de LSM.

Sensores utilizados para la captura de las señas en LSM

Como se revisó en la sección anterior, la evolución en las tecnologías de sensado ha sido notable, pudiéndose clasificar en distintas categorías basadas en la tecnología empleada y el tipo de datos que capturan. En el caso de la LSM, en estudios como el de Morfín-Chávez et al. (2023), se utilizó una cámara 2D estándar en combinación con la biblioteca MediaPipe para la segmentación de la mano y la extracción de keypoints. Este enfoque destaca por su simplicidad y accesibilidad, permitiendo la captura de imágenes de dispositivos como cámaras web o cámaras de teléfonos móviles, y la extracción de características detalladas de las manos mediante algoritmos avanzados de visión por computadora. Este método es particularmente útil para aplicaciones que requieren portabilidad, respuesta inmediata y facilidad de uso, como aplicaciones móviles o sistemas basados en la web.

Los estudios de García-Bautista et al. (2017) y Sosa-Jiménez et al. (2017), que utilizan el sensor Microsoft Kinect, destacan una categoría avanzada de dispositivos de sensado ya que está compuesto tanto de una cámara 2D tradicional como un sensor de profundidad, por lo que el sensor Kinect proporciona capacidades de seguimiento del esqueleto, lo que lo hace ideal para aplicaciones que requieren un alto nivel de detalle y precisión en el reconocimiento de gestos. Esta tecnología es capaz de capturar no solo el movimiento de las manos, sino también la postura y los

movimientos del cuerpo, ofreciendo una comprensión más completa del discurso. Es especialmente útil en entornos de investigación y aplicaciones que requieren una alta fidelidad en la captura de datos.

Retos en el reconocimiento de señas dinámicas en LSM

Las señas estáticas, que se caracterizan por la ausencia de movimiento significativo, han sido un enfoque primario en los estudios iniciales citados anteriormente. Estas señas suelen referenciar el alfabeto dactilar y los números. Mientras que las señas estáticas consisten en posturas de manos y gestos que no cambian con el tiempo, las señas dinámicas incluyen movimientos y transiciones que añaden una dimensión temporal y de fluidez al gesto, aumentando así la complejidad del reconocimiento. En el caso de las señas estáticas, los algoritmos pueden centrarse en características espaciales fijas, como la forma y orientación de la mano, lo que simplifica el proceso de identificación. En este sentido, trabajos como el de Martínez-Gutiérrez et al. (2019), Varela-Santos et al. (2021), y de Morfín-Chávez et al. (2023) se centraron en el reconocimiento de este tipo de señas, utilizando tecnologías como cámaras y software especializados para capturar y procesar las señas del alfabeto dactilar de la LSM.

Las señas dinámicas requieren un análisis detallado del movimiento de manos y otros elementos corporales, y la secuencialidad a lo largo del tiempo, lo que implica desafíos adicionales como la captura y el procesamiento de secuencias de movimiento, y la interpretación contextual de las señas. A su vez, el reconocimiento de señas puede subdividirse en reconocimiento a nivel palabra o a nivel oracional. Estudios como el de Sosa-Jiménez et al. (2022), que emplearon tecnologías como Microsoft Kinect, se han enfocado en reconocer palabras individuales en LSM, particularmente en contextos específicos como los servicios de salud. El trabajo de García-Bautista et al. (2017) utilizó Kinect y algoritmos de Distorsión Temporal Dinámica para el reconocimiento de LSM a nivel de palabra, lo que demuestra la factibilidad de capturar y procesar señas dinámicas más complejas.

El reconocimiento de LSM a nivel de oración representa un desafío mayor debido a la necesidad de comprender secuencias y contextos de gestos, lo que requiere el uso de técnicas de Procesamiento de Lenguaje

Natural (PLN) avanzadas. En este ámbito, el estudio de Mejía-Pérez et al. (2022) se destaca por utilizar redes neuronales recurrentes para procesar secuencias de señas de la LSM, abordando así el reconocimiento a niveles más avanzados.

Técnicas de IA para el reconocimiento de la LSM

El reconocimiento de la Lengua de Señas Mexicana (LSM) ha integrado diversas técnicas de inteligencia artificial (IA) para mejorar la precisión y eficiencia en la identificación de señas. A continuación se señalan cada una de las que se han aplicado en el caso de la LSM, estas abarcan las tecnologías de Redes Neuronales Artificiales (ANN), los modelos Ocultos de Markov (HMM), las Redes Neuronales Recurrentes (RNN) así como aprendizaje automático.

En el estudio de Solís et al. (2014) y de Solís et al. (2016), se implementaron Redes Neuronales Artificiales. Estas ANN se utilizaron para clasificar patrones de señas en LSM, aprovechando la capacidad de las redes para aprender y adaptarse a patrones complejos. La utilización de momentos normalizados, como características de entrada, ayudó a describir las señas de manera efectiva, permitiendo a las ANN realizar clasificaciones precisas.

Sosa-Jiménez et al. (2017) y García-Bautista et al. (2017) aplicaron Modelos Ocultos de Markov. Los HMM se usaron para modelar la secuencialidad y temporalidad de las señas en LSM, lo cual es crucial para entender gestos y movimientos que varían en el tiempo. Esta técnica es especialmente útil para capturar la dinámica y el flujo de las señas en un contexto de tiempo real.

Mejía-Pérez et al. (2023) implementaron Redes Neuronales Recurrentes con múltiples arquitecturas como LSTM (memoria a corto y largo plazo) y GRU (unidades recurrentes cerradas), desarrollando un enfoque robusto ante el ruido. Las RNN son efectivas para procesar datos secuenciales y temporales, lo cual es esencial para el reconocimiento de señas que incluyen una serie de movimientos y gestos a lo largo del tiempo. La habilidad de las RNN para manejar secuencias de datos de longitud variable las hace ideales para el análisis de LSM, que puede variar considerablemente en complejidad y duración.

Morfín-Chávez et al. (2023) aplicaron técnicas de aprendizaje automático para el reconocimiento de señas estáticas de la LSM. Se utilizaron cuatro algoritmos de clasificación diferentes: k-Nearest Neighbors (k-NN), Random Forest (RF), Naive Bayes (NB) y Support Vector Machines (SVM). Estos algoritmos fueron entrenados con los keypoints de las manos extraídos utilizando la biblioteca MediaPipe. Este estudio alcanzó resultados equiparables al uso de técnicas basadas en aprendizaje profundo.

Retos futuros hacia el reconocimiento discursivo en LSM

A lo largo de las distintas secciones de este capítulo, se ha examinado la manera en que la tecnología aplicada al reconocimiento de la Lengua de Señas Mexicana (LSM) enfrenta diversos desafíos durante sus etapas de desarrollo. En este sentido, la elección del dispositivo de sensado dependerá de las necesidades específicas del estudio o aplicación, ya que cada tecnología ofrece ventajas únicas en términos del tipo de datos capturados, el nivel de detalle y su aplicabilidad práctica. En este sentido, los desafíos abarcan desde la selección de métodos de captura hasta la gestión de la alta variabilidad entre usuarios y la similitud entre signos, lo cual puede generar ambigüedades relevantes en el discurso.

Esta variabilidad no solo surge de las diferencias individuales del habla en la ejecución de señas, sino también de las sutilezas en los gestos que pueden cambiar su significado de acuerdo con cada contexto discursivo. Esta complejidad representa un reto considerable para los sistemas de reconocimiento automático, como lo destacan los estudios de García-Bautista et al. (2017) y Mejía-Pérez et al. (2022), que buscan desarrollar algoritmos capaces de discernir estas finas diferencias. Por su parte, el estudio de Ordaz-Hernández et al. (2023) aborda este desafío mediante la creación de un corpus extenso de la LSM y el estudio de Morfín-Chávez et al. (2023) busca recopilar datos con distintos grados de variabilidad, desde señas en contextos ideales hasta señas con poca y gran variabilidad en los distintos grados de libertad de la mano.

Respecto a las condiciones de captura, Solís et al. (2016) y Solís et al. (2014) se enfocaron en el reconocimiento de LSM bajo diferentes condiciones de iluminación. En este tenor, fue crucial integrar tecnologías de

video capaces de identificar y rastrear en 3D la configuración manual así como los rasgos no manuales, mediante el uso de hardware como cámaras de profundidad o estéreo 3D y a través de software que estimara la profundidad mediante las librerías OpenPose o MediaPipe. Estos datos de alto nivel pueden ser aprovechados por técnicas que tomen series discursivas temporales como input de entrada, por ejemplo, aplicables para los HMM (Sosa-Jiménez et al. 2022) y las RNN (Mejía-Pérez et al. 2022).

La identificación de señas en un contexto dinámico presenta desafíos adicionales debido a la naturaleza fluida y compleja de la comunicación en LSM. A diferencia de otros idiomas donde las palabras pueden ser identificadas en un contexto estático y lineal, las lenguas de modalidad visogestual incluyen dos articuladores manuales, los cuales pueden ofrecer información independiente cada uno de estos. Además, las señas de la LSM pueden variar significativamente de acuerdo con la velocidad del movimiento, la posición de las manos y el movimiento del cuerpo puede codificar aspectos discursivos como fuente de información. Asimismo, los rasgos nomanuales codifican rasgos de aspectualidad y modalidad. Esto hace que el proceso de reconocimiento sea más complejo y requiera algoritmos y modelos avanzados capaces de capturar estas variaciones dinámicas de manera precisa.

En este sentido, la comunicación en una lengua de modalidad visogestual es intrínsecamente dinámica y depende del flujo continuo e interacción entre gestualidad y expresión lingüística. De igual manera, para que un sistema de reconocimiento sea efectivo y funcional en un contexto cotidiano, debe ser capaz de procesar e interpretar señas a medida que ocurren, sin retrasos significativos. Esto es esencial no solo para capturar el contenido léxico de las señas, sino también para preservar los elementos prosódicos y aquellos fuera del dominio manual como las expresiones faciales y el movimiento del cuerpo los cuales son fundamentales para el significado y la estructura gramatical en las lenguas de señas.

En el caso de la LSM, la falta de bancos de datos glosados específicos representa un obstáculo significativo para el desarrollo de sistemas de reconocimiento automático. Estos bancos de datos son fundamentales para entrenar y validar modelos de reconocimiento de señas, ya que proporcionan ítems etiquetados que permiten a los algoritmos aprender y mejorar

su desempeño. Sin suficientes datos glosados disponibles, el desarrollo de sistemas de reconocimiento de LSM se ve limitado en términos de precisión y capacidad de generalización.

Otro desafío en la identificación de señas en LSM es la codificación de configuraciones manuales específicas que no están incluidas en el alfabeto manual estándar. Estas configuraciones manuales, pueden incluir clasificadores de forma y de tamaño que codifican detalles sutiles en la forma de los objetos a los que hacen referencia. Sin embargo una representación estándar para estas configuraciones manuales puede dificultar la tarea de identificación de señas, ya que los algoritmos deben ser capaces de reconocer y distinguir una amplia variedad de formas y movimientos de manos.

La eficiencia en tiempo real es otro aspecto vital para garantizar una interacción fluida y natural entre los usuarios de una lengua de señas y un sistema interactivo o en situaciones comunicativas entre un usuario nativo de la LSM y personas que no lo son, por ejemplo, puede ser aplicable en situaciones donde la comunicación oportuna es crítica, como en la educación, el ámbito laboral, y en servicios de emergencia y atención médica. La capacidad de convertir las señas en texto o habla en tiempo real permite una comunicación bidireccional efectiva y accesible. Además, un sistema que funcione en tiempo real tiene aplicaciones potenciales en la educación de la lengua de señas con usuarios que desean aprender una lengua visogestual como L2, proporcionando retroalimentación instantánea a los estudiantes. Esto no solo aumenta la accesibilidad para las personas sordas o con problemas de audición, sino que también facilita la inclusión y la igualdad en diversos ámbitos de la vida social y profesional. Estudios como los de García-Bautista et al. (2017), Sosa-Jimenez et al. (2022), y Mejía-Pérez et al. (2022) han hecho importantes avances en esta dirección e ilustran la diversidad y complejidad de los desafíos enfrentados en el reconocimiento automático de LSM, así como los esfuerzos continuos para superar estas barreras mediante el desarrollo y mejora de tecnologías de inteligencia artificial en LSM.

Las futuras investigaciones en el reconocimiento automático de la LSM pueden abordarse desde varias perspectivas. A partir de los estudios recientes, se han identificado tres áreas clave para el desarrollo futuro:

La necesidad de mejorar la etapa de captura de datos en el reconocimiento automático de la LSM es evidente en diversos estudios recientes.

En primer lugar, la variabilidad en la ejecución de señas presenta un desafío significativo que requiere una mayor diversidad y tamaño en la población de muestra para una representación más completa y precisa (Galicia et al., 2015; Sosa-Jiménez et al., 2022). Además, el uso de diferentes tecnologías de sensado ha demostrado ser fundamental para capturar con precisión la complejidad de las señas, como se ve en los trabajos de García-Bautista et al. (2017) y Mejía-Pérez et al. (2022). Estas investigaciones resaltan la importancia de incorporar una variedad de métodos de sensado y un espectro amplio de usuarios para mejorar la precisión y eficacia en el reconocimiento, abordando tanto la diversidad de configuraciones manuales como las diferencias en la realización de las señas así como las diferencias del habla entre los usuarios (Rios-Figueroa et al., 2022; Morfín-Chávez et al., 2023). Por otro lado, se recomienda que los modelos de reconocimiento debieran ser independientes de la tecnología existente, de modo que puedan adaptarse a nuevos y más avanzados sensores.

El desarrollo de sistemas de reconocimiento de la LSM adaptados a contextos específicos, como el educativo, médico o laboral, representa un área de investigación emergente y crucial para la inclusión y accesibilidad. El estudio de Sosa-Jimenez et al. (2022) es un ejemplo de cómo los sistemas de reconocimiento de LSM pueden ser adaptados para contextos de atención médica. Por otro lado, el trabajo de Gortarez-Pelayo et al. (2023) está diseñado para que la población que no conoce la LSM se familiarice con el alfabeto manual fomentando así la cultura de la lengua de señas. Trabajos futuros podrían enfocarse en adaptar y personalizar los sistemas de reconocimiento para otros entornos específicos, mejorando así su utilidad y aplicabilidad.

La innovación en nuevos métodos y algoritmos de IA para el reconocimiento de la LSM es una necesidad imperante, tal como se evidencia en investigaciones recientes. Los estudios actuales, como los realizados por García-Bautista et al. (2017) y Mejía-Pérez et al. (2022), han demostrado avances significativos utilizando métodos de aprendizaje profundo potentes. Sin embargo, para abordar de manera efectiva la complejidad y diversidad de la LSM, es esencial desarrollar métodos más sofisticados y algoritmos avanzados. Esto incluye la exploración de nuevas técnicas de IA no solo para mejorar la precisión, sino también la velocidad en la interpretación de las señas, especialmente en la detección conjunta de gestos

manuales sutiles y expresiones faciales. Asimismo, trabajos como el de Ríos-Figueroa et al. (2022) y Morfín-Chávez et al. (2023) indican la dirección hacia la que deben orientarse las investigaciones futuras, enfatizando la importancia de la innovación en el procesamiento de los datos, desde el reconocimiento de señas estáticas hasta la traducción de frases que incluya la estructura gramatical del español, así como aprovechar el análisis contextual para asistir en el proceso de reconocimiento.

Conclusiones

La investigación y desarrollo en el reconocimiento automático de lenguas de señas son clave para una sociedad más inclusiva y accesible, particularmente para la comunidad sorda. Avances como el perfeccionamiento de redes neuronales y el procesamiento de imágenes están rompiendo barreras lingüísticas y culturales, permitiendo una comunicación fluida y equitativa entre personas sordas y oyentes. En este sentido, el continuo desarrollo e investigación en materia de reconocimiento automático de lenguas de modalidad visogestual desempeña un papel fundamental en la construcción de conocimiento acerca de la codificación lingüística de una lengua de modalidad visogestual y es esencial para conocer más profundamente la gramática de la LSM, conocimiento especialmente útil para el desarrollo de aplicaciones diseñadas para la comunidad sorda.

Actualmente, los avances significativos, como el refinamiento de las redes neuronales y el procesamiento de imágenes, están derribando barreras lingüísticas, facilitando una comunicación fluida y equitativa entre personas sordas y oyentes, pero también tendrán aplicaciones en materia educativa tanto para la comunidad sorda como para oyentes. En este sentido, las mejoras tecnológicas en la actualidad no solo aumentarán la precisión y la adaptabilidad de las herramientas de comunicación, sino que también crearán en pocos años nuevas oportunidades en ámbitos educativos, laborales y sociales para las personas sordas. El progreso en este campo demuestra un compromiso con la igualdad, impulsando innovaciones que benefician a una amplia gama de aplicaciones y mejoran la calidad de vida e interacción social para la comunidad silente.

Es importante destacar que las investigaciones en materia de recono-

cimiento de patrones lingüísticos de lenguas visogestuales no solo representan avances tecnológicos, sino también un paso crucial hacia un futuro más empático y equitativo. Un futuro donde todos los individuos puedan comunicarse y contribuir sin restricciones, fortaleciendo así los valores fundamentales de igualdad y respeto en nuestra sociedad. Este trabajo no solo abre nuevas posibilidades de comunicación, sino que también promueve un cambio cultural significativo ante el desarrollo cada vez más complejo en materia tecnológica y del procesamiento de información humana.

Como hemos mencionado, estas mejoras tecnológicas no solo aumentan la precisión y adaptabilidad de las herramientas de comunicación, sino que también abren oportunidades en ámbitos educativos, laborales y sociales para las personas sordas. En este sentido, el progreso en este campo refleja un compromiso con la diversidad y la igualdad, impulsando innovaciones que benefician un amplio rango de aplicaciones y potencialmente pueden mejorar la calidad de vida de sus usuarios. Este tipo de investigaciones, más allá de ser un avance tecnológico, es un paso crucial hacia un futuro más empático y equitativo, donde todos los individuos puedan comunicarse y contribuir sin restricciones, fortaleciendo los valores de igualdad y respeto en nuestra sociedad.

Referencias

- Albanie, S. et al. (2020). BSL-1K: Scaling Up Co-articulated Sign Language Recognition Using Mouthing Cues. En Vedaldi, A., Bischof, H., Brox, T., Frahm, J. M. (Eds.), *Computer Vision – ECCV 2020. ECCV 2020. Lecture Notes in Computer Science* (Vol. 12356). Springer, Cham. https://doi.org/10.1007/978-3-030-58621-8_3
- Aly, S. y Aly, W. (2020). *DeepArSLR: A novel signer-independent deep learning framework for isolated Arabic sign language gestures recognition. IEEE access: practical innovations, open solutions*, 8, 83199-83212. <https://doi.org/10.1109/access.2020.2990699>
- Agarwal, P., Liao, J., Hooper, S., & Sperling, R. (2021). *Applying deep learning to a sign language progress monitoring system. Distance Learning*, 18(3), 45-55. [https://eric.ed.gov/?q=artificial+intelligence+AND+sign+language+&ff1=subArtificial+Intelligence&ff2=dtySince_2020&id=EJ1327302]
- Bragg, D., Koller, O., Bellard, M., Berke, L., Boudreault, P., Braffort, A., Caselli, N., Huenerfauth, M., Kacorri, H., Verhoef, T., Vogler, C. y Ringel Morris, M. (2019). *Sign language recognition, generation, and translation: An interdisciplinary perspective. The 21st International ACM SIGACCESS Conference on Computers and Accessibility.*
- Camgoz, N.C.; Hadfield, S.; Koller, O.; Ney, H.; Bowden, R. (2018). Neural sign language translation. En *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA.*
- Camgoz, N. C., Hadfield, S., Koller, O. y Bowden, R. (2017). *SubUNets: End-to-end Hand Shape and Continuous Sign Language Recognition. *ICCV.**
- Elatawy, S. M., Hawa, D. M., Ewees, A. A. y Saad, A. M. (2020). *Recognition system for alphabet Arabic sign language using neutrosophic and fuzzy c-means. Education and Information Technologies*, 25(6), 5601–5616. <https://doi.org/10.1007/s10639-020-10184-6>
- Forster, J.; Schmidt, C.; Koller, O.; Bellgardt, M.; Ney, H. (2014). Extensions of the Sign Language Recognition and Translation Corpus RWTH-PHOENIX-Weather. En *Proceedings of the LREC*, (pp. 1911–1916). Reykjavik, Iceland.

- Fridman-Mintz, B. (2005). *LSM tense & aspect inflections. A Dissertation Submitted to the Faculty of the Graduate School of Arts and Sciences of Georgetown University*. https://www.academia.edu/59937562/LSM_tense_and_aspect_inflections
- Galicia, R., Carranza, O., Jiménez, E. D, Rivera, G. E. (2015). *Mexican sign language recognition using movement sensor* (pp. 573-578).
- García-Bautista, G., Trujillo-Romero, F., y Caballero-Morales, S. (2017). *Mexican sign language recognition using kinect and data time warping algorithm*. 2017 International Conference On Electronics, Communications And Computers (CONIELECOMP). <https://doi.org/10.1109/conielecomp.2017.7891832>
- Gellai, D. B. (2023). Enterprising academics: Heterarchical policy networks for artificial intelligence in British higher education. *ECNU Review of Education*, 6(4), 568–596. [<http://files.eric.ed.gov/fulltext/EJ1401707.pdf>]
- Gortarez-Pelayo, J. J., Morfín-Chávez, R. F., & Lopez-Nava, I. H. (2023, noviembre). DAKTILOS: An Interactive Platform for Teaching Mexican Sign Language (LSM). En *International Conference on Ubiquitous Computing and Ambient Intelligence* (pp. 264-269). Cham: Springer Nature Switzerland.
- Herrera Crespo, J. (2021). *Diseño y validación de un software de reconocimiento de gestos sirviéndose del acelerómetro de un reloj inteligente*. Universitat Politècnica de València. <http://hdl.handle.net/10251/172272>
- INALSA (s/f). *Proyecto de Ley de Reconocimiento de la Lengua de Señas Argentina - LSA, y de creación del Instituto Nacional de Lengua de Señas Argentina*. [<https://inalsa.cas.org.ar/nuestra-lsa/la-ls-en-el-mundo/#:~:text=Ethnologue%2C%20una%20enciclopedia%20que%20cataloga,señas%20dispersas%20alrededor%20del%20mundo.>]
- Kothadiya, D., Bhatt, C., Sapariya, K., Patel, K., Gil, A., & Corchado Rodríguez, J. (2022). *Deepsign: Sign Language Detection and Recognition Using Deep Learning*. *Electronics*, 11, 1780. <https://doi.org/10.3390/electronics11111780>
- Koller, O., Zargaran, S. y Ney, H. (2015). *Continuous sign language recognition: Towards large vocabulary statistical recognition systems handling multiple signers*. *Computer Vision and Image Understanding*, 141, 108-125.

- Lang, H., Manganelli, J., & Panchanathan, S. (2020). *Virtual reality and augmented reality in social learning spaces for people with visual impairments*. *IEEE Transactions on Learning Technologies*, 13(2), 422-433.
- Liang, X., Angelopoulou, A., Kapetanios, E., Woll, B., Al Batat, R., & Woolfe, T. (2020). A multi-modal machine learning approach and toolkit to automate recognition of early stages of dementia among british sign language users. En *Computer Vision—ECCV 2020 Workshops: Glasgow, UK, August 23–28, 2020, Proceedings, Part II* 16 (pp. 278-293). Springer International Publishing.
- Martínez-Gutiérrez, M. E., Rojano-Cáceres, J. R., Benítez-Guerrero, E., y Sánchez-Barrera, H. E. (2019). *Data acquisition software for sign language recognition*. *Res. Comput. Sci.*, 148(3), 205–211
- Mejía-Peréz, K., Córdova-Esparza, D.-M., Terven, J., Herrera-Navarro, A.-M., García-Ramírez, T. y Ramírez-Pedraza, A. (2022). *Automatic recognition of Mexican Sign Language using a depth camera and recurrent neural networks*. *Applied Sciences (Basel, Switzerland)*, 12(11), 5523. <https://doi.org/10.3390/app12115523>
- Moher, David, Shamseer, Larissa, Clarke, Mike, Gherzi, Davina, Liberati, Alessandro, Petticrew, Mark, Shekelle, Paul, & Stewart, Lesley A. (2016). Ítems de referencia para publicar Protocolos de Revisiones Sistemáticas y Metaanálisis: declaración PRISMA-P 2015. *Revista Española de Nutrición Humana y Dietética*, 20(2), 148-160. <https://dx.doi.org/10.14306/renhyd.20.2.223>
- Morfín-Chávez, R. F., Gortarez-Pelayo, J. J., & Lopez-Nava, I. H. (2023). Fingerspelling Recognition in Mexican Sign Language (LSM) Using Machine Learning. En *Mexican International Conference on Artificial Intelligence* (pp. 110-120). Cham: Springer Nature Switzerland.
- Murakami, K., & Taguchi, H. (1991). *Gesture recognition using recurrent neural networks*. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.
- Ordaz-Hernández, K., Castillo-Gaytán, D., Rodríguez-Recio, A. S., Boone-Obregón, R. D., Hernández-García, L. Á., & Hilario-Acuapan, G. (2023). *Hacia la construcción de una base de datos abierta de la LSM*. *Pädi Boletín Científico de Ciencias Básicas e Ingenierías del ICBI*, 11. (Especial 2), 134-141.

- Ompi. (2019). *BrightSign: el guante inteligente que da voz a quienes no pueden hablar*. (s. f). Wipo.int. Recuperado el 5 de febrero de 2024, de https://www.wipo.int/wipo_magazine/es/2019/05/article_0005.html
- Papastratis, I., Chatzikonstantinou, C., Konstantinidis, D., Dimitropoulos, K., & Daras, P. (2021). *Artificial intelligence technologies for sign language*. *Sensors* (Basel, Switzerland), 21(17), 5843. <https://doi.org/10.3390/s21175843>
- Pigou, L., Dieleman, S., Kindermans, P. J., & Schrauwen, B. (2015). *Sign language recognition using convolutional neural networks*. *Computer Vision and Pattern Recognition Workshops*.
- Quandt, L. (2020). *Teaching ASL Signs using Signing Avatars and Immersive Learning in Virtual Reality*. The 22nd International ACM SIGACCESS Conference on Computers and Accessibility. doi:10.1145/3373625.3418042
- Rios-Figueroa, H. V., Sánchez-García, A. J., Sosa-Jiménez, C. O., & Solís-González-Cosío, A. L. (2022). *Use of spherical and Cartesian features for learning and recognition of the static Mexican Sign Language alphabet*. *Mathematics*, 10(16), 2904. <https://doi.org/10.3390/math10162904>
- Schioppo Zachary Meyer Deigo Fabiano Shaun Canavan. (2019). Learning Sign Language in a Virtual Environment. En *Proceedings of ACM CHI conference*, Glasgow, UK, May 2019 (CHI '19), 6 pages. https://doi.org/10.475/123_4
- Solís, F. J., Toxqui, C., & Martínez, D. (2015). *Mexican Sign Language Recognition Using Jacobi-Fourier Moments*.
- Solis, F. J., Martínez, D., & Espinoza, O. (2016). *Automatic Mexican Sign Language Recognition Using Normalized Moments and Artificial Neural Networks*.
- Solís-V., J.-F., Toxqui-Quitl, C., Martínez-Martínez, D. y Margarita H.-G. (2014). *Mexican sign language recognition using normalized moments and artificial neural networks*.
- Sosa-Jimenez, C., Rios-Figueroa, H., Rechy-Ramirez, E., Marin-Hernandez, A., y Gonzalez-Cosio, A. (2017). *Real-time Mexican Sign Language recognition*. *IEEE International Autumn Meeting On Power, Electronics And Computing (ROPEC)*. <https://doi.org/10.1109/ropec.2017.8261606>
- Sosa-Jimenez, C. O., Rios-Figueroa, H., & Ana Luisa Solís-González-Cosío. (2022). *A Prototype for Mexican Sign Language Recognition and Synthesis in Support of a Primary Care Physician*.

- Sturman, D. J., & Zeltzer, D. (1994). A Survey of Glove-Based Input. *IEEE Computer Graphics and Applications*, 14(1), 30-39.
- Sreemathy, R., Turuk, M., Kulkarni, I. y Khurana, S. (2023). *Sign language recognition using artificial intelligence*. *Education and Information Technologies*, 28(5), 5259–5278. (<https://doi.org/10.1007/s10639-022-11391-z>]
- Sugandhi, P. Kumar y Kaur, S. (2021). Indian Sign Language Generation System. En *Computer*, 54(3), 37-46. doi: 10.1109/MC.2020.2992237.
- Varela-Santos, H., Morales-Jiménez, A., Córdova-Esparza, D. M., Terven, J., Mirelez-Delgado, F. D. y Orenday-Delgado, A. (2021). *Assistive device for the translation from mexican sign language to verbal language*. *Computación y Sistemas*, 25(3), 451-464.